DOCUMENT RESUME

ED 127 373 95 TM 005 533

AUTHOR Elias, Patricia J.; Hare, Gail
TITLE Beginning Teacher Evaluation Study: Phase II, 1973-74, Final Report: Volume V.4. Special Study: Reading Difficulties: Viedetape Test of Teachers' Diagnostic Skills.
INSTITUTION California State Commission for Teacher Preparation and Licensing, Sacramento.; Educational Testing Service, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ETS-PR-76-15
PUB DATE Jul 76
NOTE 105p.; For related documents, see TM 005 524-535

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS Criterion Referenced Tests; *Diagnostic Tests; Effective Teaching; Elementary Education; *Elementary School Teachers; *Films; Grade 2; Grade 5; *Reading Diagnosis; *Reading Difficulty; Teacher Behavior; Teaching Skills; Video Tape Recordings

ABSTRACT

The Beginning Teacher Evaluation Study (BTES), Phase II, was a research project on effective teaching behavior--what teachers do that significantly affects what and how pupils learn. The purposes of Phase II were to (1) develop an assessment system for measuring teacher and pupil behaviors and other factors which could influence each of them and their interrelationships and (2) generate hypotheses about the interrelationships among teacher and pupil behaviors and related factors. Forty-one second grade and 54 fifth grade experienced teachers participated in the study. In this volume of the final report, teacher performance on the Diagnostic Film Test was analytically related to teacher performance on other reading and reading related tests, to observed teacher responses to pupils in their classrooms, and to observed and measured pupil reading performances. The results are discussed. The Diagnostic Reading Test was standardized, utilized a real life situation, and was criterion referenced. It was felt that the instrument had potential as both a training and selection tool and was worthy of further research and development. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED127373

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
DOROTHY URBAN
ETS
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

TM005 533

PR-76-15

BEGINNING TEACHER EVALUATION STUDY
PHASE II 1973-74

FINAL REPORT

VOLUME V.4.

SPECIAL STUDY

READING DIFFICULTIES: VIDEOTAPE TEST OF TEACHERS' DIAGNOSTIC SKILLS

BY

PATRICIA ELIAS
GAIL HARE

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

BEGINNING TEACHER EVALUATION STUDY: PHASE II

FINAL REPORT: VOLUME V.4.

SPECIAL STUDY

READING DIFFICULTIES: VIDEOTAPE TEST OF TEACHERS' DIAGNOSTIC SKILLS

by

Patricia Elias

Gail Hare

Educational Testing Service
Berkeley, California

A project conducted by Educational Testing Service for the California Commission for Teacher Preparation and Licensing and funded by the National Institute of Education.

Copyright © 1975, 1976 by Educational Testing Service. All rights reserved.

PREFACE

The Beginning Teacher Evaluation Study (BTES) is a long-term project of the California Commission for Teacher Preparation and Licensing. The Commission is responsible for licensing teachers in California and is trying to determine what factors should be considered in this process.

The second phase of the study was conducted by Educational Testing Service for the Commission. Phase II was the hypotheses-generating and instrument-development phase of BTES. ETS had two tasks: (1) to develop an assessment system to measure both teacher and pupil behaviors as well as other factors which might be related to these behaviors; and (2) to generate hypotheses about the interrelationships between teacher and pupil behaviors and related factors.

The study was conducted in 43 schools in eight districts throughout the state of California. A total of 41 second grade teachers and 54 fifth grade teachers participated in the project during Phase II.

The final report for Phase II consists of several volumes. Volume I describes the design and rationale for the experimental design and data analysis procedures and includes the major findings of Phase II. Volume II describes the conduct of the field study and the sample of participants.

Because of the complex nature of Phase II, a variety of techniques was used to measure teacher and pupil behaviors. They are described in Volumes III, IV, and V. Results are also included in these volumes.

Volume III describes the observation systems in detail and is available in three separately bound sections. The first section, Volume III.1., describes the behavior recording observation system used in the project--APPLE (Anectdotal Process for Promoting the Learning Experience). Volume III.2. describes the category system used to observe classroom activities--

RAMOS (Reading and Mathematics Observation System). The third section of this volume, III.3., covers the videotaping of instructional activities during reading and mathematics.

Volume IV concerns other aspects of the measurement system and covers both the pupil and teacher test batteries.

The fifth volume covers a series of small studies done as part of Phase II. Volume V.1. looks at teacher aptitudes as related to teacher behaviors. Volume V.2. is concerned with the relationship between teacher expectations and pupil performance. Volume V.3. reviews performance of pupils in the BTES teachers' classrooms for two years prior to Phase II, the historical test data. Volume V.4. discusses the Diagnostic Film Test, a device designed to assess teachers' skills in diagnosing reading problems and prescribing corrective action. Volume V.5. summarizes the results of work diaries completed by the teachers on their reading and mathematics instructional program.

Information on the availability of these volumes can be obtained from:

Dr. Frederick J. McDonald
Educational Studies
Educational Testing Service
Princeton, NJ 08540

Information on other phases of BTES can be obtained from:

California Commission for Teacher
Preparation and Licensing
1020 O Street
Sacramento, CA 95814

ACKNOWLEDGEMENTS

The authors would like to acknowledge the work of Reginald Corder, Ann Piestrup, Pamela Roelfs, Mary Corder and Helga Grabske. The study could not have been completed without their efforts. We particularly want to thank the teachers who contributed their time and thoughts.

Patricia Elias
Gail Hare

TABLE OF CONTENTS

	Page
Preface	i
Acknowledgements	iii
The Diagnostic Film Test	1
Background	1
Test Development	2
Error types selected for inclusion	4
Selection of materials	5
Test format	5
Selection of the Sample and Administration of the Test	7
Scoring the Test	9
Results	12
Indicators of the Validity of the Diagnostic Film Test	22
Teacher Performance on Reading and Reading Related Tests and Diagnostic Skills	22
Observed Incidence of Specific Kinds of Teaching Performance and Teacher Diagnostic Skills	28
Observed Incidence of Specific Kinds of Positive Pupil Learning and Teacher Diagnostic Skills	34
Teacher Diagnostic Skills and Pupil Skill Growth in Reading	38
Summary and Conclusions	44
References	47
Appendix A: Reading Difficulties: Videotape Test of Teachers' Diagnostic Skills	49
Appendix B: Diagnostic Film Test Scoring Manual	75

LIST OF TABLES

Table		Page
1	Mean Scores of Second Grade Teachers on Subparts and on the Total Test	13
2	Mean Scores of Fifth Grade Teachers on Subparts and on the Total Test	14
3	Mean Scores of the Total Teacher Sample on Subparts and on the Total Test	15
4	Correlation Coefficients Between the Subpart Scores of the Second Grade Teacher Sample.	18
5	Correlation Coefficients Between the Subpart Scores of the Fifth Grade Teacher Sample	19
6	Correlation Coefficients Between the Subpart Scores of the Total Teacher Sample	20
7	Correlations Between Second Grade Teacher Scores on Reading and Related Tests and Scores on the Diagnostic Film Test	26
8	Correlations Between Fifth Grade Teacher Scores on Reading and Related Tests and Scores on the Diagnostic Film Test	27
9	Correlations Between Observed Teacher Responses During Reading and Scores on the Diagnostic Film Test: Second Grade.	32
10	Correlations Between Observed Teacher Responses During Reading and Scores on the Diagnostic Film Test: Fifth Grade	33
11	Correlations Between Observed Indications of Pupil Reading Skills and Scores on the Diagnostic Film Test: Second Grade	35
12	Correlations Between Observed Indications of Pupil Reading Skills and Scores on the Diagnostic Film Test: Fifth Grade.	37
13	Correlation Coefficients Between Second Grade Teacher Scores on the Diagnostic Film Test and Spring Pupil Test Scores Partialling Out Fall Test Scores.	40
14	Correlation Coefficients Between Fifth Grade Teacher Scores on the Diagnostic Film Test and Spring Pupil Test Scores Partialling Out Fall Test Scores.	41

LIST OF FIGURES

Figure		Page
1	Pupil assignments to presentations	3
2	The assignment of comprehension errors to pupils	6
3	Numbers of teachers selected for the Diagnostic Film Test sample	8
4	A model of the hypothetical relationships among teacher diagnostic skills, other teacher characteristics, and pupil behavior	23
5	APPLE lexicon definition of selected teacher responses	29

The Diagnostic Film Test

One of the special studies for Phase II of the Beginning Teacher Evaluation Study was to develop a videotaped test of teachers' skills in diagnosing the reading difficulties of children. The test was a series of episodes in which children are depicted reading. Each child demonstrated some difficulty in performing a particular reading skill. The reading skills included were similar to those identified in the BTES reading matrix and measured in the BTES reading test battery for pupils. Second and fifth grade teachers were asked to diagnose each child's problem and to prescribe an appropriate strategy or strategies for correcting the problem.

The proposed requisites of such a test for assessing teachers' skills in diagnosing reading problems were:

1. It should be standardized, so any idiosyncratic teacher-student relationships would not confound the diagnosis.
2. It should utilize a real life situation, employing the same types of information a teacher is typically required to use in diagnosing reading problems.
3. It should be criterion referenced, in the sense that a "good" teacher should be able to diagnose each reading difficulty portrayed.

Background

To prepare for the development of the Diagnostic Film Test, the ETS staff* reviewed relevant literature from research studies on the

---

*The ETS staff was augmented by Ann McCormick Piestrup, whose areas of expertise included reading and linguistics. Her skills and experience included the production and utilization of videotapes as educational tools. Her work was essential to the development and field testing of the Diagnostic Film Test.

types and frequencies of oral reading errors made by elementary school pupils. It had been our experience that teachers of young pupils spend a fair amount of classroom time listening to individual children read aloud. It was something of a surprise, therefore, to discover the relative paucity of publications about the kinds of errors associated with oral reading, in contrast to the amount of research available on the errors children make on paper-and-pencil tests. The literature that was reviewed is listed in the references section on page 47.

The review of the literature served primarily to document the second and third of the three prerequisites listed above: to enable us to present situations in which children were depicted making the kinds of errors that young readers really make; and to allow the presentation of vignettes in which the criterion errors were represented in proportion to the degree that they occur in oral reading. The assumption was that the most frequently occurring oral reading errors identified by researchers who worked in classrooms as well as clinical settings would be those that a teacher would need to be aware of in order to diagnose reading problems.

Test Development

Two types of content, word lists and paragraphs, were included in the test. Four presentations of each type of content were developed to represent with some fidelity the common types and proportional frequency of the errors identified in previous research. To control or "allow for", in the experimental sense, some of the possible idiosyncratic teacher-pupil relationships mentioned in the first of the three prerequisites, a specific pattern of pupil assignment to each of the eight

presentations was planned. Equal numbers of boys and girls, blacks and whites, were included among the eight children. The assignment of children to each type of material is presented in Figure 1.

Pupil Assignment to Word Lists

	Black	White
Male	W.L. III	W.L. IV
Female	W.L. I	W.L. II

Pupil Assignment to Paragraphs

	Black	White
Male	Para. IV	Para. III
Female	Para. I	Para. II

Figure 1. Pupil assignments to presentations.

To meet the first objective, standardization in the classic measurement sense was also built into the test; the same materials were presented and the same administrative procedures were established for use in all testing sessions. A third and perhaps more subtle form of standardization was also imposed. The test reflects our bias that elementary school teachers at all levels are challenged by the need to help children at all stages in their reading development, children whose skills range from the pre-primer level to fairly sophisticated levels of reading ability. Therefore, we concentrated on developing a single test which would begin to reflect the range of oral reading problems represented in elementary school classrooms.

To keep the content of the presentations as logically and psychometrically separate as possible, materials were developed which featured each type of criterion error in single filmed vignettes. Each word list was used to reflect a criterion decoding error; each paragraph was used to reflect a criterion comprehension-related error. Four dominant error types were selected for each of the two kinds of

content. The developmental procedures were similar for the word lists and the paragraphs. The developmental sequence involved first, the review of the research literature on oral reading errors, noting types of mistakes. It proceeded through categorization of the errors, and focused on those which decline with age or proficiency. The next major step was to select four "major" types of errors and "distractor" or "minor" errors for each type of content (word list or paragraph). The final steps in the developmental sequence were to review related tests and diagnostic materials and to select materials for inclusion in the eight vignettes of the Diagnostic Film Test.

Error types selected for inclusion. After review and categorization, the following types of "major" and "distractor" or "minor" errors were selected for inclusion in the Diagnostic Film Test.

I. "Major" Types of Paragraph Errors

A. Omission of words, letters
B. Substitution of graphically similar words that do not make sense
C. Substitution of phonologically similar words that do not make sense
D. Substitution of words that are phonologically and graphically similar that do make sense

II. "Minor" Types of Paragraph Errors

Note: These are less serious errors in the context of reading considered as a progressive learning process in which children approximate words using graphic clues, meaning context and grammatical context.

A. Hesitation
B. Repetition
C. Insertion

III. "Major" Types of Word Decoding Errors

A. Initial word sound errors, primarily consonant related
B. Final word sound errors, primarily consonant related
C. Medial word sound errors, primarily vowel related
D. Medial word sound errors, primarily consonant related

IV. "Minor" Types of Word Decoding Errors

A. Letter sequence reversal errors*
B. Single examples of each of two types of "major" errors, complementary to the type emphasized in the word list

In each of the word list readings, six examples of the "major" error are presented, and there are three single occurance "minor" errors. Four examples of each of the "major" errors occurred in each of the paragraph readings. The three "minor" errors for the paragraph readings occurred once per paragraph. Following each reading of a paragraph, the child was asked four questions, two of which involved literal comprehension and two of which involved inferential reasoning. Either the two literal questions or the two inferential questions were answered incorrectly by the child reading the paragraph.

Selection of materials. There were several considerations in establishing the criteria for the selection of the materials which would form the content of the test. For both the paragraphs and the word lists, the materials selected had to be of an appropriate difficulty level and of known (graded) difficulty level across the four vignettes within each content area. Several other criteria also obtained for the paragraph materials. The length of paragraphs had to be similar and appropriate (60 to 80 words). To match the BTES Phase II comprehension matrix, two kinds of questions associated with the paragraphs were needed: content or literal meaning versus inference or interpretation.

Test format. The word lists were developed; the paragraphs and related inferential and literal questions were selected from the

---

*This type of decoding error occurred less than fifteen percent of the time when children read word lists designed so that each of the "major" errors listed previously had an equal opportunity to occur. (Shankweiler and Liberman, 1972.)

Cooperative Primary Test battery (Educational Testing Service, 1965-67). One further concern about the influence on teachers of idiosyncratic pupil-teacher relationships was obviated by assigning the two types of comprehension errors to specific types of children. Literal interpretation is a less demanding comprehension skill than is the ability to make inferences. The assignment of those two criterion errors was balanced in the following pattern.

	Black		White	
	Male	Female	Male	Female
Inference	Correct Answers	Incorrect Answers	Incorrect Answers	Correct Answers
Literal	Incorrect Answers	Correct Answers	Correct Answers	Incorrect Answers

Figure 2. The assignment of comprehension errors to pupils.

The printed materials teachers would use during the Diagnostic Film Test were developed and prepared. They included three parts: Section One accompanied the films of children reading word lists; Section Two went with the reading of the paragraphs; and Section Three was to be presented with a replay of the fourth and fifth taped vignettes. Each written section included a correct version of the materials read in each filmed presentation for teacher use. The response mode built into Sections One and Two was open ended questions. They were:

1. What is *(child's name)*'s reading problem or problems? Be as specific as you can.
2. What strategies or approaches would you use to help *(child's name)* improve his/her reading skills?
3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?
   Materials:
   Human:
4. What resources would you need (if any) that are not presently available in your school or district?
   Materials:
   Human:

The presentations were designed so that when teachers had viewed the eight reading sequences and responded in Sections One and Two of the printed materials, those materials were collected. The videotape was then rewound to present the last of the word lists and the first of the paragraph reading film segments. Teachers were given Section Three and were asked to respond to written multiple choice or recognition format questions about the two presentations. The "major" and the "minor" error categories were included in the multiple choice lists. The multiple choice format was similar to that used in paper-and-pencil tests of teachers' knowledge. The Diagnostic Film Test was designed to more closely approximate the kinds of experiences teachers have while teaching reading than do paper-and-pencil tests of diagnostic skills. We were interested in the differences, if any, between knowledge and application of that knowledge. These printed materials are included in Appendix A.

Selection of the Sample and Administration of the Test

The objective for the Diagnostic Film Test, during this year of instrument development and the generation of hypotheses, was to pilot test the instrument on a sample of Phase II teachers. The sample was

chosen in a specific way. We wanted a sample that included a fairly equal number of second and fifth grade teachers to allow a reasonable pilot test of the Diagnostic Film Test. We wanted to maximize the opportunity to generate a variety of hypotheses about teacher attributes that might or might not impinge on how teachers conceptualize oral reading performance. The staff generated lists of the Phase II teachers who, during the fall testing sessions, had:

1. accurately ranked the reading performance of their students (about 95 percent accuracy);
2. inaccurately ranked the reading performance (about 40 percent accuracy);
3. accurately ranked for some groups of children and not for others, e.g., boys versus girls (about 70 percent total accuracy).

A sample of 35 teachers was chosen which included 18 at the second grade level and 17 at the fifth grade level. The sample included eight teachers who ranked accurately; 12 teachers who were inaccurate in their rankings; and 15 teachers with mixed ranking ability. The sample is represented below.

Ranking Ability	Grade Level Taught: Second	Grade Level Taught: Fifth
Accurate	5	3
Mixed	6	9
Inaccurate	7	5

Figure 3. Numbers of teachers selected for the Diagnostic Film Test sample.

The Diagnostic Film Test was administered to the selected sample of teachers as part of their spring testing battery. The directions given to the teachers were very general. We wanted to approximate the classroom setting where the teacher has to decide whether a child has a reading problem based on decoding errors or on comprehension errors. A Sony portable videorecorder and a Sony monitor were used. After the first presentation session, the tape was rewound and shown a second time. Teachers appeared to watch the second presentation in order to confirm their diagnoses.

Scoring the Test

Scoring procedures were developed for each of the three sections of the Diagnostic Film Test. Basic guidelines were established for socring the test prior to the test administration. Actual Phase II teacher responses were recorded which indicated skills in the diagnosis and remediation of the reading difficulties presented by each of the eight children in the taped vignettes. Their responses were used to illustrate the scoring models for the test.

The scoring standards were fairly rigorous. Teachers were required to identify the major difficulty depicted and to suggest some strategy or strategies for correcting that problem. Credit was also given for the identification of the minor problems depicted and their remediation strategies. Teachers were given credit for any and all strategies they suggested which were associated with a major reading problem. They were credited with only one strategy for each minor error they identified, no matter how many they provided. We did not impose our notion of what a "good" remediation strategy might be. The strategy

merely had to be associated with an error in the teacher's test booklet to be scored.

It was not required that teachers respond in exact "reading specialist" terminology. For example, if a teacher indicated that a child was having problems with the sounds in the middle of words such as t or p, that was accepted as an equivalent response to the identification of the child's problem as medial consonant errors. The teachers occasionally supplied diagnostic information in the remediation sections of the test. Acceptable responses were scored no matter where teachers happened to record them within the test booklet.

The maximum possible score for the diagnosis of the reading problems in Section One and in Section Two was seven points for each of the eight presentations. The maximum possible score for the strategy listings of Section One and Section Two is not known, but it is not likely to exceed thirteen points. It is related to the number of strategies a teacher might list for each major or dominant error type. Section Three, the multiple choice format, afforded a total possible score of fourteen points. The complete scoring manual is included in Appendix B.

The tests were handscored by a research assistant trained in the system. They were independently rescored by another trained research assistant to determine the inter-rater reliability of the scoring system. The inter-rater reliability of the two was .997* which was so high that we questioned the results. All identification was then removed from about one-half of the protocols and they were

---

*Spearman Rank Order Correlation Coefficient.

rescored by the original scorer. The results were about the same. Both research assistants had worked with the test authors on the development of the scoring manual, and had selected the illustrative examples for the manual. It is probable that in more general field use with less highly trained personnel the inter-rater reliability would be somewhat lower.

Results

The scores of teachers were organized to reflect scores on each part of the test, totals on various combinations, and the total test. The average score on the free response decoding questions (Section One of the test) is shown separately and in combination with the results of the decoding question in multiple choice format (Section Three of the test). The comprehension scores are reflected in the same way. Total scores on the free response questions from the first two sections of the test and total scores on Section Three, the multiple choice or recognition tasks, are also shown. One of the areas we wanted to explore was the relationship between teachers' responses to open ended questions and to questions with a recognition format. The multiple choice format is the most frequently used convention in tests of diagnostic skills.

In general, the means of the second grade teachers were higher than those of the fifth grade teachers (Tables 1 and 2). On the free response decoding sections of the test in particular, the second grade teachers were more able to accurately diagnose and suggest strategies for remediation of oral reading problems than their fifth grade colleagues. The scores of the two groups on the open ended comprehension sections were more similar. The scores of the two groups on the multiple choice questions were very much alike. Scores for the entire sample are shown in Table 3.

It was theoretically possible to score at least 80 points on Section One, the decoding free response problems. The number of strategies which a teacher might present to remediate a major problem

Table 1

Mean Scores of Second Grade Teachers on Subparts and on the Total Test

Diagnostic Film Test Scoring Categories	Teacher Scores Mean	Standard Deviation	Number
Decoding: Free Response	21.41	9.26	17
Comprehension: Free Response	11.69	7.65	16
Decoding: Multiple Choice	3.94	1.27	17
Comprehension: Multiple Choice	3.15	1.48	17
Decoding: Total Score	25.35	9.41	17
Comprehension: Total Score	14.84	8.26	16
Free Response Questions: Total	33.31	15.24	16
Multiple Choice Questions: Total	7.06	2.01	17
Test Total Score	40.38	15.86	16

Table 2

Mean Scores of Fifth Grade Teachers on Subparts and on the Total Test

Diagnostic Film Test Scoring Categories	Teacher Scores Mean	Standard Deviation	Number
Decoding: Free Response	13.47	8.64	17
Comprehension: Free Response	10.88	6.53	16
Decoding: Multiple Choice	3.35	1.47	17
Comprehension: Multiple Choice	3.21	1.44	17
Decoding: Total Score	17.41	9.33	17
Comprehension: Total Score	14.13	6.70	16
Free Response Questions: Total	24.38	13.31	16
Multiple Choice Questions: Total	6.56	2.51	17
Test Total Score	31.22	14.34	16

Table 3

Mean Scores of the Total
Teacher Sample on Subparts and on
the Total Test

Diagnostic Film Test Scoring Categories	Teacher Scores Mean	Standard Deviation	Number
Decoding: Free Response	17.44	9.70	34
Comprehension: Free Response	11.28	7.01	32
Decoding: Multiple Choice	3.65	1.39	34
Comprehension: Multiple Choice	3.18	1.44	34
Decoding: Total Score	21.38	10.07	34
Comprehension: Total Score	14.48	7.40	32
Free Response Questions: Total	28.84	14.79	32
Multiple Choice Questions: Total	6.81	2.25	34
Test Total Score	35.80	15.58	32

was not limited. The eighty point theoretical score could be achieved if five strategies were presented for each of the major errors in the four cases in Section One. The average scores actually obtained by second and fifth grade teachers were quite low (21.41 and 13.47 respectively). The range of scores was from ten to 42 for second grade teachers and from zero to 27 for fifth grade teachers.

Section Two, the comprehension free response problems, also had a potential total score of 80. The average score of second grade teachers was 11.69 and of fifth grade teachers, 10.88. The range was from two to 23 for second grade teachers and from three to 26 for fifth grade teachers. Again teacher performance was low in relationship to the possible points obtainable.

Section Three had a possible total score of fourteen points. A comprehension case and a decoding case were included and the teacher was required to check the child's problem or problems and to circle the major problem or problems on a list. Both groups of teachers did better, in terms of the maximum possible score, on the multiple choice questions than they did on open ended questions. The average second grade teacher score was 7.06 and that of fifth grade teachers was 6.56. The range for second grade teachers was from 4.0 to 11.0, and for fifth, from 1.5 to 10.5.

The total test score range was from 18.5 to 71.0 for second grade teachers and from 8.5 to 59.5 for fifth grade teachers. A non-parametric analog to the parametric t test, the Mann-Whitney U Test (Siegel, 1956), was used to examine the significance, if any, of the differences between scores of second and fifth grade teachers. We

used one-tailed tests to examine the hypotheses that the scores of second grade teachers on decoding problems and on the total test were significantly better than those of fifth grade teachers, and that the comprehension scores of fifth grade teachers were significantly better than those of second grade teachers. The scores of second grade teachers on the open ended decoding problems were significantly higher than those of fifth grade teachers ($p \leq .05$) as were the total test scores ($p \leq .05$); but, the scores of the two groups of teachers on the open ended comprehension problems were not statistically different. The statistical tests verified what was apparent in the tables.

Correlations among the subparts of the test for second, fifth and the total sample of teachers are shown in Tables 4, 5, and 6. For second and fifth grade teachers, the correlations between Section One (decoding-free response) and Section Two (comprehension-free response) suggest that the two parts of the test are related but measure somewhat different kinds of skills (.57 and .49 respectively).

Scores on the open ended decoding questions were not related to scores on the multiple choice decoding problem for either second or fifth grade teachers, nor were scores on the open ended comprehension problems related to scores on the multiple choice comprehension questions. The multiple choice decoding problem was not related to the total test scores of either group, but the comprehension questions in that format did correlate with total test scores. This suggests that the teacher scores on the multiple choice questions were contributing little or nothing to teacher performance on the test. Although they have been included as scoring categories in later tabular presentations in the report, the discussion will focus on

Table 4

Correlation Coefficients Between the Subpart Scores of the Second Grade Teacher Sample

Diagnostic Film Test Scoring Categories	Decoding (F.R.)	Comprehension (F.R.)	Decoding (M.C.)	Comprehension (M.C.)	Decoding (Total)	Comprehension (Total)	Free Response Questions (Total)	Multiple Choice Questions (Total)	Test Total
Decoding: Free Response	--								
Comprehension: Free Response	.57(.011)**	--							
Decoding: Multiple Choice	N.S.*	N.S.	--						
Comprehension: Multiple Choice	N.S.	N.S.	N.S.	--					
Decoding: Total Score	.99(.001)	.57(.011)	N.S.	N.S.	--				
Comprehension: Total Score	.56(.012)	.99(.001)	N.S.	.48(.031)	.56(.013)	--			
Free Response Questions: Total	.91(.001)	.86(.001)	N.S.	N.S.	.90(.001)	.84(.001)	--		
Multiple Choice Questions: Total	N.S.	N.S.	.69(.002)	.77(.001)	N.S.	.38(.073)	N.S.	--	
Test Total Score	.90(.001)	.86(.001)	N.S.	.36(.088)	.90(.001)	.86(.001)	.99(.001)	.35(.093)	--

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Table 5

Correlation Coefficients Between the Subpart Scores of the Fifth Grade Teacher Sample

Diagnostic Film Test Scoring Categories	Decoding (F.R.)	Comprehension (F.R.)	Decoding (M.C.)	Comprehension (M.C.)	Decoding (Total)	Comprehension (Total)	Free Response Questions (Total)	Multiple Choice Questions (Total)	Test Total
Decoding: Free Response	--								
Comprehension: Free Response	.49(.027)**	--							
Decoding: Multiple Choice	N.S.*	N.S.	--						
Comprehension: Multiple Choice	.51(.018)	N.S.	.50(.022)	--					
Decoding: Total Score	.97(.001)	.43(.050)	N.S.	.54(.013)	--				
Comprehension: Total Score	.59(.009)	.98(.001)	N.S.	N.S.	.54(.017)	--			
Free Response Questions: Total	.90(.001)	.81(.001)	N.S.	N.S.	.85(.001)	.87(.001)	--		
Multiple Choice Questions: Total	.43(.043)	N.S.	.87(.001)	.86(.001)	.48(.026)	N.S.	N.S.	--	
Test Total Score	.92(.001)	.77(.001)	N.S.	.47(.035)	.88(.001)	.85(.001)	.98(.001)	.46(.037)	--

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Table 6

Correlation Coefficients Between the Subpart Scores of the Total Teacher Sample

Diagnostic Film Test Scoring Categories	Decoding (F.R.)	Comprehension (F.R.)	Decoding (M.C.)	Comprehension (M.C.)	Decoding (Total)	Comprehension (Total)	Free Response Questions (Total)	Multiple Choice Questions (Total)	Test Total
Decoding: Free Response	--								
Comprehension: Free Response	.51(.002)**	--							
Decoding: Multiple Choice	N.S.*	N.S.	--						
Comprehension: Multiple Choice	.30(.046)	N.S.	.28(.052)	--					
Decoding: Total Score	.98(.001)	.49(.003)	.30(.043)	.32(.033)	--				
Comprehension: Total Score	.54(.001)	.98(.001)	N.S.	.36(.022)	.52(.002)	--			
Free Response Questions: Total	.91(.001)	.82(.001)	N.S.	.27(.068)	.89(.001)	.83(.001)	--		
Multiple Choice Questions: Total	.32(.035)	N.S.	.80(.001)	.81(.001)	.38(.013)	.31(.045)	.27(.071)	--	
Test Total Score	.91(.001)	.80(.001)	.70(.068)	.38(.017)	.90(.001)	.83(.001)	.99(.001)	.41(.011)	--

* Not significant ($\rho > .10$)
** Levels of significance indicated in parentheses

the open ended questions. The relatively better performance of teachers on the multiple choice questions, and the lack of correlational relationships with free responses at least tentatively suggests that teacher performance may be different on the two types of test formats. Further research is warranted to test the effects of format on teacher demonstrations of diagnostic skills. That research might include verification of test results in the classroom, where, presumably, teacher performance is more dependent on the kind of recall which is assessed by the open ended questions. Effective classroom teachers are able to generate accurate diagnoses, not just to recognize them. There appears to be a distinction between knowledge of diagnostic skills and the application of that knowledge.

The sample of teachers for the Diagnostic Film Test was selected to include teachers with varying abilities to rank expected pupil performance. The scores of the teachers with good, mixed and poor ranking skills were compared at each grade level. A Kruskal-Wallis One-Way Analysis of Variance by ranks (Siegel, 1956), a nonparametric test, was run to test the hypothesis that the film test scores of the three groups of teachers were not statistically different. The test results revealed no significant relationships between the teachers' ability to judge a pupil's performance in reading and their diagnostic skills. A teacher may know whether or not a child will be successful without knowing why or what to do about it.

Indicators of the Validity of the Diagnostic Film Test

The Diagnostic Film Test was one of several approaches to describing teacher performance in the project. Teacher performance was also assessed by paper-and-pencil tests of aptitude and knowledge and by the direct observation of actual teaching. The validity of the Diagnostic Film Test as a measure of teachers' ability to diagnose oral reading problems was explored in comparison to teacher performance on selected paper-and-pencil measures, to observed teacher performance, to observed pupil performance, and to pupil reading achievement. Teacher scores on the Diagnostic Film Test was related to:

1) teachers' knowledge about teaching reading;
2) teachers' attitudes about teaching reading;
3) teachers' knowledge of teaching another academic subject (mathematics);
4) observed teacher performance during reading instruction;
5) observed evidence of positive pupil performance during reading instructional activities; and,
6) pupil achievement in reading.

The relationships in the model we are testing are shown in Figure 4.

Teacher Performance on Reading and Reading Related Tests and Diagnostic Skills

The correlations between teacher scores on paper-and-pencil tests of their knowledge about teaching reading, their ability to apply knowledge to reading problems, their attitude toward teaching reading and the various score configurations from the Diagnostic Film Test are

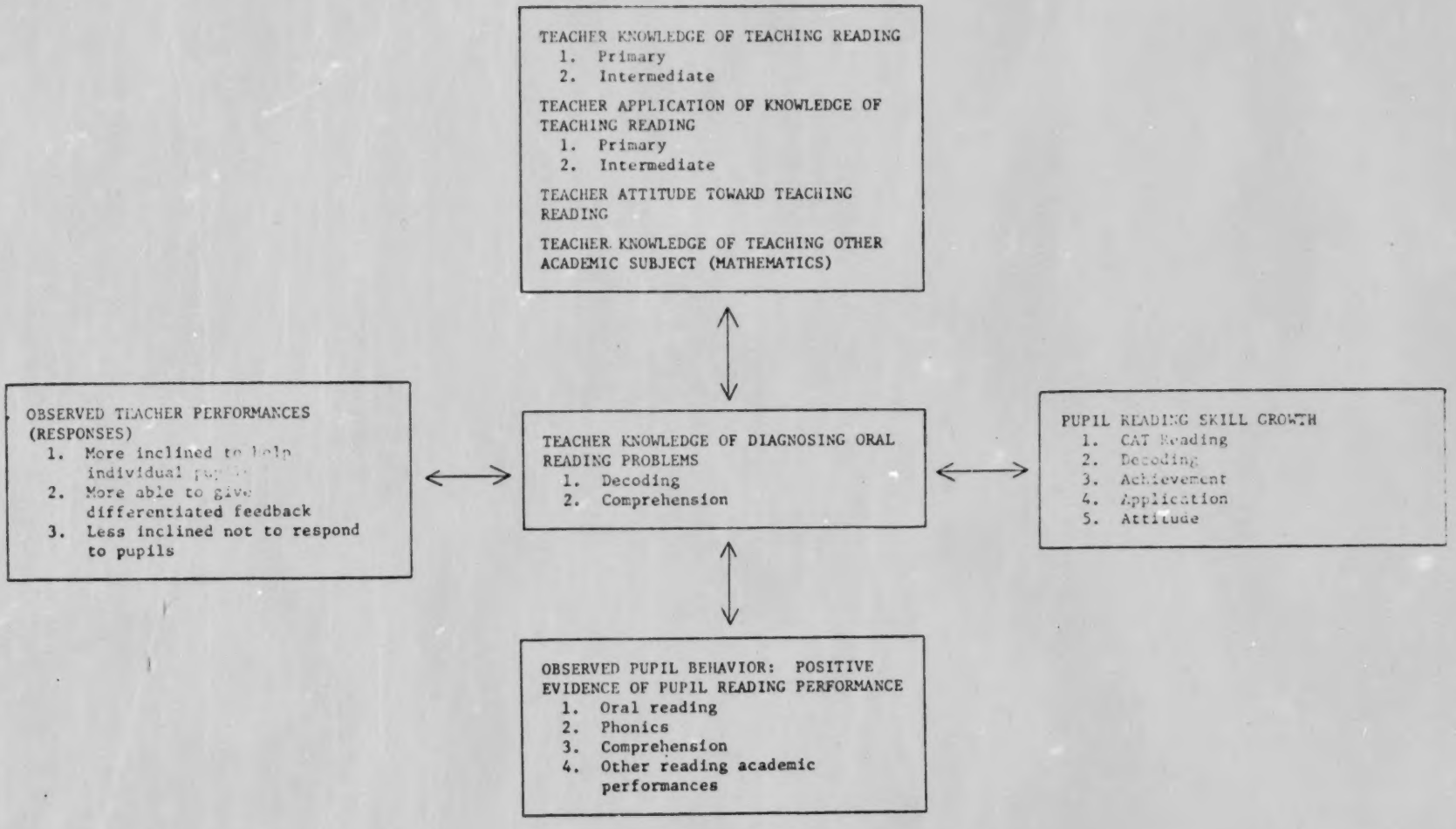
TEACHER KNOWLEDGE OF TEACHING READING
1. Primary
2. Intermediate
TEACHER APPLICATION OF KNOWLEDGE OF TEACHING READING
1. Primary
2. Intermediate
TEACHER ATTITUDE TOWARD TEACHING READING
TEACHER KNOWLEDGE OF TEACHING OTHER ACADEMIC SUBJECT (MATHEMATICS)
OBSERVED TEACHER PERFORMANCES (RESPONSES)
1. More inclined to help individual pupils
2. More able to give differentiated feedback
3. Less inclined not to respond to pupils
TEACHER KNOWLEDGE OF DIAGNOSING ORAL READING PROBLEMS
1. Decoding
2. Comprehension
PUPIL READING SKILL GROWTH
1. CAT Reading
2. Decoding
3. Achievement
4. Application
5. Attitude
OBSERVED PUPIL BEHAVIOR: POSITIVE EVIDENCE OF PUPIL READING PERFORMANCE
1. Oral reading
2. Phonics
3. Comprehension
4. Other reading academic performances

Figure 4. A model of the hypothetical relationships among teacher diagnostic skills, other teacher characteristics, and pupil behavior.

shown in Tables 7 and 8. Second grade teachers completed eight item tests of Primary Knowledge and Primary Application. Fifth grade teachers completed similar tests of Intermediate Knowledge and Intermediate Application. Examples of questions from the Primary Knowledge test are:

Primary children who read fog for frog, and gas for glass have not yet mastered

A. vowel digraphs
B. compound words
C. initial consonants
D. consonant blends
E. double consonants

What are the "Dolch Words"?

A. The first 100 words in the "Dick and Jane" basal reader.
B. Any list of words which a teacher uses to test phonics knowledge.
C. A list of 220 very common words which includes most function words.
D. A list of words for beginning language experience reading.
E. A list of words used in traffic signs.

In the Application tests, teachers were asked to indicate for each of five responses, which might be reasonable teacher actions, which actions would not help, and which actions were questionable for each of the problems posed. Questions for teachers included:

A beginning reader reads help for home when he reads aloud.

_____A. Have him use the words correctly in an oral sentence

_____B. Correct him every time he makes this mistake.

_____C. Find out whether he knows the correspondences between letters and sounds for long and short vowels.

_____D. Have him write each word several times.

_____E. Show him the different configurations and help him to learn that home has more "short letters" than help.

An intermediate girl has trouble with words like addition and dictionary because of the way she tries to pronounce the t in these words.

_____A. She is a sight reader and has learned the wrong pronunciation patterns. You must correct her every time she mispronounces these words.

_____B. She does not know that tion says "shun" and needs practice in reading words with this spelling.

_____C. She has a basic reading problem that is just beginning to show up in the middle grades. She needs tutoring from the Reading Specialist.

_____D. She is not a good speller and should copy the words and say the letters softly to herself.

_____E. She does not use context in reading or she would be able to figure these words out. Tell her to try another guess each time she mispronounces them.

It was hypothesized that the content of these short tests would relate to teachers' ability to diagnose oral reading problems. There were no relationships between the knowledge and application of knowledge tests and any major scoring category on the film test for second grade teachers. We considered the three free-response categories, decoding, comprehension, and their total, to be major categories of scores. The only second grade correlation was a negative one between application skills and the one multiple choice question on the film test about decoding problems (Table 7). For fifth grade teachers, however, consistent correlations emerged between knowledge and teachers' diagnostic skills (Table 8). Fifth grade teachers who had better scores on the paper-and-pencil test of general knowledge about teaching reading had better scores on all of the major scoring configurations of the film test. There was only one (out of six) significant relationship between teachers' positive attitudes about

Table 7

Correlations Between Second Grade Teacher Scores on Reading and Related Tests and Scores on the Diagnostic Film Test

Diagnostic Film Test Scoring Categories	Teacher Tests			
	Primary Knowledge	Primary Application	Attitude: Reading	Mathematics Teaching Knowledge
Decoding: Free Response	NS*	NS	NS	.36(.077)**
Comprehension: Free Response	NS	NS	NS	NS
Free Response Questions: Total	NS	NS	NS	.37(.079)
Decoding: Multiple Choice	NS	-.49(.025)	NS	NS
Comprehension: Multiple Choice	NS	NS	NS	NS
Decoding: Total Score	NS	NS	NS	.38(.067)
Comprehension: Total Score	NS	NS	NS	NS
Multiple Choice Questions: Total	NS	NS	NS	NS
Test Total Score	NS	NS	NS	.41(.058)

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Table 8

Correlations Between Fifth Grade Teacher Scores on Reading and Related Tests and Scores on the Diagnostic Film Test

Diagnostic Film Test Scoring Categories	Teacher Tests			
	Intermediate Knowledge	Intermediate Application	Attitude: Reading	Mathematics Teaching Knowledge
Decoding: Free Response	.48(.026)**	NS*	.37(.072)	.33(.098)
Comprehension: Free Response	.50(.025)	NS	NS	NS
Free Response Questions: Total	.62(.006)	NS	NS	NS
Decoding: Multiple Choice	NS	NS	NS	NS
Comprehension: Multiple Choice	.34(.089)	NS	.36(.082)	.49(.023)
Decoding: Total Score	.38(.065)	NS	.41(.051)	NS
Comprehension: Total Score	.61(.006)	.36(.086)	NS	NS
Multiple Choice Questions: Total	.36(.077)	NS	NS	.47(.028)
Test Total Score	.56(.013)	NS	NS	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses.

teaching reading and a major scoring category. Fifth grade teachers with positive attitudes about reading knew more about the diagnosis and remediation of decoding problems.

It was hypothesized, for the model being tested in this section of the report, that teachers who know more about one academic area, the diagnosis of reading problems, would know more about the teaching of the other major elementary school academic subject, mathematics. That hypothesized relationship was supported at both grade levels, but in a specific way. The fifth and second grade teachers who knew more about the diagnosis and remediation of decoding problems also knew more about the teaching of mathematics.

There are at least tentative indications that teacher performance on the Diagnostic Film Test related to other paper-and-pencil tests of general knowledge of teaching reading and mathematics. The relationships were more frequent for fifth than second grade teachers. Performance on the film test does not appear to relate to performance on paper-and-pencil tests of attitudes about reading or to the application of knowledge to reading problems.

Observed Incidence of Specific Kinds of Teaching Performance and Teacher Diagnostic Skills

One of the two observation systems utilized to gather Phase II data was APPLE* (Anecdotal Processing to Promote the Learning Experience), a system in which observers generated a low inference record of teacher and pupil behavior. Among the things observers recorded was a continuous

---

*APPLE was developed by Nadine Lambert of the University of California, Berkeley. It is described in its application in this project in Volume III.1.

INSTRUCTING:	Teacher activities connected with her traditional jobs of instructing and supervising the everyday business of the classroom -- "gave a lesson," directed, organized, "called on pupil," answered a question.
DISCIPLINING:	Punishment imposed on child for his behavior, e.g. sent to principal's office or told to sit in corner or to stay after school.
TEACHER HELP:	Teacher gives some specific instructional assistance to an individual student.
IGNORING:	Teacher seemingly aware but not responding to behavior or performance of pupil.
MOVES ON:	Teacher does not comment on or give feedback to a pupil response, as in a question and answer session; instead, she calls on another pupil or moves on to another part of the lesson. Can occur when pupil's response is correct or incorrect.
NEGATIVE FEEDBACK:	Mild negative verbal reinforcement in response to academic or behavioral performance (e.g., "wrong," "no").
NONE:	No response to pupil performance or behavior. No distinction made between whether teacher aware of this performance or behavior or not (unless is clearly a situation of IGNORING).
POSITIVE FEEDBACK:	Use of mild positive verbal reinforcement in response to academic or behavioral performance (e.g., "right," "O.K.," "good," "fine," "correct").
PRAISE:	Teacher commends or tangibly rewards a student's activities or products.
QUESTIONING:	Teacher responds to pupil's behavior or answer to a previous question by herself asking a question.
RECOGNIZING:	Teacher's neutral nonevaluative response to an academic performance or a nonacademic behavior (e.g., "notices,", "acknowledges").
REDIRECTING:	A verbal or non-verbal action by the teacher to redirect a pupil's behavior. Done without apparent anger.

Figure 5. APPLE lexicon definition of selected teacher responses.

Copyright © 1971, 1976 by N. M. Lambert, et. al. All rights reserved.

record of teacher responses to the behavior of selected target pupils. Responses of teachers were then categorized. The APPLE lexicon general definition of the teacher responses is shown in Figure 5. For the analysis discussed in this section, only the observer description of teacher responses to pupils during reading instructional activities were included.

Eight target pupils* were identified for observation in each classroom. Pupils and appropriate alternates were selected by comparing teacher rankings of expected pupil performance with the actual fall reading performance on one or more of the Phase II tests. A boy and a girl were selected from the top, just above the middle, just below the middle and from the bottom of the distributions in each classroom. The list was also balanced to reflect the ethnic composition in each classroom. The analyses in this and the next section of the report compare the observed activities of teachers in response to a sample of less able, average and above average pupils in each classroom.

We hypothesized that teachers who knew more about the diagnosis and remediation of oral reading problems would be those who would be more inclined to help individual pupils (because they were more able to provide specific kinds of assistance); those who would be less inclined <u>not</u> to respond to individual pupil performance (more awareness of a pupil's problems might make it more difficult to ignore them); and, those who would give more differentiated feedback,

---

*The selection process for target pupils is described more fully in Volume II.

positive and negative, to pupil performance (teachers with more diagnostic skills would respond to correct and incorrect pupil performance more consistently). The APPLE Teacher Response categories selected to test these hypotheses were "Teacher Help," "None," "Positive Feedback," and "Negative Feedback."

Second grade teachers who were frequently observed to be giving some specific instructional assistance to an individual student during reading ("Teacher Help") were those with better diagnostic skills, and were not the teachers who were observed to make no response to pupil performance or behavior, the APPLE category "None" (Table 9). The size and levels of significance of the correlations between "Teacher Help" and diagnostic skills of second grade teachers are extraordinary (from .67 to .81). That relationship should be tested in future research and by further examination of the Phase II data base to determine what kinds of target pupils the second grade teachers are working with.

The correlations in Table 9 suggest only modest support for the notion of the association of differentiated feedback and the diagnostic and remediation skills of second grade teachers. That relationship appears more consistent for fifth grade teachers, at least in terms of "Positive Feedback" (Table 10). It appears that there is some support for the hypothesis that teachers, particularly second grade teachers, who are better able to diagnose oral reading problems and to generate remediation strategies respond more to individual students in their classrooms during reading activities than do teachers with less well developed diagnostic and remediation skills.

Table 9

Correlations Between Observed Teacher Responses During Reading and Scores on the Diagnostic Film Test: Second Grade

Diagnostic Film Test Scoring Categories	Teacher Responses			
	Teacher Help	None	Positive Feedback	Negative Feedback
Decoding: Free Response	.75(.001)**	-.38(.066)	NS*	NS
Comprehension: Free Response	.67(.003)	NS	NS	.42(.052)
Free Response Questions: Total	.81(.001)	-.41(.056)	NS	NS
Decoding: Multiple Choice	NS	.36(.078)	NS	NS
Comprehension: Multiple Choice	NS	NS	NS	-.36(.077)
Decoding: Total Score	.72(.001)	NS	NS	NS
Comprehension: Total Score	.64(.004)	NS	NS	NS
Multiple Choice Questions: Total	NS	NS	NS	NS
Test Total Score	.79(.001)	-.38(.075)	NS	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Table 10

Correlations Between Observed Teacher Responses During Reading and Scores on the Diagnostic Film Test: Fifth Grade

Diagnostic Film Test Scoring Categories	Teacher Responses			
	Teacher Help	None	Positive Feedback	Negative Feedback
Decoding: Free Response	NS*	NS	.34(.091)**	NS
Comprehension: Free Response	NS	NS	NS	NS
Free Response Questions: Total	NS	NS	.37(.079)	NS
Decoding: Multiple Choice	NS	NS	NS	NS
Comprehension: Multiple Choice	NS	NS	NS	NS
Decoding: Total Score	NS	NS	.38(.066)	NS
Comprehension: Total Score	NS	NS	NS	NS
Multiple Choice Questions: Total	NS	NS	NS	NS
Test Total Score	NS	NS	.36(.085)	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Observed Incidence of Specific Kinds of Positive Pupil Learning and Teacher Diagnostic Skills

We hypothesized that the pupils of teachers with more diagnostic skills would be observed to be doing well in oral reading, phonics, comprehension, and other academic reading tasks. To test that hypothesis, observed evidence of positive pupil performance in those kinds of activities was compared to teacher performance on the Diagnostic Film Test (Tables 11 and 12).

The APPLE lexicon definition of a "+ Phonics Skills" pupil event, shown in the table headings, is "positive evidence of skill in sounding out new words; knowing the sound connected with letter." If observers had any doubt about the quality of a pupil activity, they were trained to rate it as a "0" or neutral event. (The rating system was identical for all pupil events; a "+", "0", or "-" was assigned to the written description of a pupil's activity.) Similarly "+ Oral Reading" and "+ Comprehension" were positive evidence of pupil skills in those activities. All other kinds of positive reading performances were categorized as "+ Academic Performance."

The diagnostic skills of the second grade teachers were consistently and positively associated with observed evidence of positive pupil performance on phonics tasks, and had some relationship to observed good general reading performance (Table 11). However, good teacher performance on the film test was consistently and negatively associated with good pupil performance on comprehension tasks in second grade.

That conundrum could relate to the priorities of the second grade teachers with more highly developed diagnostic and remediation skills.

Table 11

Correlations Between Observed Indications of Pupil Reading Skills and Scores on the Diagnostic Film Test: Second Grade

Diagnostic Film Test Scoring Categories	Pupil Events			
	Oral Reading	Phonics Skills	Comprehension	Academic Performance
Decoding: Free Response	NS*	.35(.087)**	-.41(.051)	NS
Comprehension: Free Response	NS	.43(.050)	-.38(.075)	.39(.086)
Free Response Questions: Total	NS	.48(.029)	-.38(.074)	NS
Decoding: Multiple Choice	NS	NS	NS	NS
Comprehension: Multiple Choice	-.53(.015)	NS	NS	.37(.075)
Decoding: Total Score	NS	NS	-.36(.080)	NS
Comprehension: Total Score	NS	.43(.047)	NS	.43(.047)
Multiple Choice Questions: Total	-.34(.094)	NS	NS	NS
Test Total Score	NS	.45(.042)	-.35(.093)	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

They perhaps concentrate on the development of decoding skills to the exclusion of skill development in comprehension. It is also possible that the more diagnostically skilled teachers could be expending much of their teaching efforts on less able pupils, which at the second grade level would include pupils with poorly developed phonics or decoding skills.

It is important to note that in their responses to the film test, very few second or fifth grade teachers attended to the nature of the errors presented by the videotaped pupil responses to literal and inferential comprehension questions. Teachers responded to the more specific word attack and decoding problems evidenced during the videotaped reading of paragraphs. That is, even during the sections of the test focused on comprehension problems, teachers noticed word related errors rather than the more sophisticated reading problems inherent in depicted lack of literal or inferential comprehension skills.

The relationships which emerged for grade five teachers are consistent with the hypothesis that teacher diagnostic skills relate to observed pupil reading skills (Table 12). Fifth grade teachers who know more about the diagnosis and remediation of decoding problems had pupils who were observed to be doing well in oral reading, comprehension and general reading kinds of tasks. However, there were no significant relationships with observed indications of pupil skills in phonics. It would be interesting to examine the observational records of the less able target pupils in these classrooms to determine in what reading activities such pupils were involved and to ascertain what materials they were using. It seems probable that some less able fifth grade pupils would need phonics skills instruction.

Table 12

Correlations Between Observed Indications of Pupil Reading Skills and Scores on the Diagnostic Film Test: Fifth Grade

Diagnostic Film Test Scoring Categories	Pupil Events			
	Oral Reading	Phonics Skills	Comprehension	Academic Performance
Decoding: Free Response	.35(.087)**	NS*	.45(.036)	.36(.081)
Comprehension: Free Response	NS	NS	.35(.090)	NS
Free Response Questions: Total	NS	NS	.41(.058)	.36(.084)
Decoding: Multiple Choice	NS	NS	NS	NS
Comprehension: Multiple Choice	NS	NS	NS	NS
Decoding: Total Score	NS	NS	.50(.021)	.43(.042)
Comprehension: Total Score	NS	NS	.34(.099)	NS
Multiple Choice Questions: Total	NS	NS	NS	NS
Test Total Score	.37(.082)	NS	.44(.045)	.38(.077)

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Teacher Diagnostic Skills and Pupil Skill Growth in Reading

The indicators of the validity of the Diagnostic Film Test that will be explored in this section of the report are whether teacher diagnostic and remediation skills related to growth in pupil performance as measured by the Phase II pupil test battery.* We were interested in the relationship of teacher film test scores to pupil performance in the spring adjusted for the same pupil's performance in the fall on the criterion variables in reading. The statistical technique used, partial correlation, allows the researcher to calculate a measure of association between two variables while adjusting for the effects of one or more additional or control variables.

Measures of reading achievement, comprehension, application, and decoding skills were obtained for pupils during the fall and in the spring of the study. Pupil scores on each of those measures and an attitude toward reading measure controlled for pupil entry level skills are the criterion measures to which teacher film test scores have been compared. All pupils for whom fall and spring scores were obtained on any single criterion measure were included in the analysis.

The results are fascinating (Tables 13 and 14). So far in this section we have attempted to delineate the relationships among several kinds of teacher or pupil information as indicators of the validity of the Diagnostic Film Test as a measure of a presumably valuable set of teaching skills. We have offered support for the generalization that the diagnostic and remediation skills of teachers is related to their

---

*The Phase II pupil test battery is described in detail in Volume IV.

knowledge about teaching reading; and, we have provided information that linked teacher diagnostic and remediation skills to both observed teacher responses to pupils and to observed evidence of pupil skills in reading. This may be one of the few instances in reading research of demonstrated connections between performance on a paper-and-pencil test of teaching skill and observed pupil behavior and teacher behavior in the classroom.

However, on the last of the validity indicators to be tested for the hypothetical model we used--the relationship of pupil skill growth as measured by some of the paper-and-pencil tests of reading skills to the diagnostic and remediation skills of their teachers--the relationships are consistently negative (Tables 13 and 14). The only exception among the major scoring configurations was a positive relationship between the attitude of second grade pupils toward reading and the test score category of teachers' skills on comprehension problems.

The pupils of second grade teachers with stronger diagnostic and remediation skills evidenced less skill growth on the measures of reading achievement and application skills than did pupils of less skilled teachers (Table 13). The pupils of fifth grade teachers with more skills at diagnosing and suggesting remediation strategies for comprehension problems did less well in skill growth on the reading achievement test (Table 14). Pupil skill growth on measures of decoding was not related to teacher diagnostic skills at either grade level.

All pupils in a classroom were used as the unit of analysis for this report. It was suggested previously that teachers with highly developed diagnostic skills may concentrate their teaching energies

Table 13

Correlation Coefficients Between Second Grade Teacher Scores on the Diagnostic Film Test and Spring Pupil Test Scores Partialling Out Fall Test Scores

Diagnostic Film Test Scoring Categories	Pupil Reading Tests				
	CAT Reading	Achievement	Decoding	Application	Attitude
Decoding: Free Response	NS*	-.45(.062)**	NS	-.40(.091)	NS
Comprehension: Free Response	NS	-.50(.040)	NS	NS	.43(.071)
Free Response Questions: Total	NS	-.57(.023)	NS	-.41(.080)	NS
Decoding: Multiple Choice	NS	NS	NS	.48(.050)	NS
Comprehension: Multiple Choice	.54(.029)	NS	NS	NS	NS
Decoding: Total Score	NS	-.39(.095)	NS	NS	NS
Comprehension: Total Score	NS	-.54(.029)	NS	NS	NS
Multiple Choice Questions: Total	.38(.100)	NS	NS	.39(.094)	NS
Test Total Score	NS	-.55(.28)	NS	NS	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

Table 14

Correlation Coefficients Between Fifth Grade Teacher Scores on the Diagnostic Film Test and Spring Pupil Test Scores Partialling Out Fall Test Scores

Diagnostic Film Test Scoring Categories	Pupil Reading Tests				
	CAT Reading	Achievement	Decoding	Application	Attitude
Decoding: Free Response	NS*	NS	NS	NS	NS
Comprehension: Free Response	NS	-.43(.055)**	NS	NS	NS
Free Response Questions: Total	NS	NS	NS	NS	NS
Decoding: Multiple Choice	-.39(.074)	NS	NS	-.47(.039)	NS
Comprehension: Multiple Choice	NS	NS	NS	NS	NS
Decoding: Total Score	NS	NS	NS	-.38(.081)	NS
Comprehension: Total Score	NS	-.47(.039)	NS	NS	NS
Multiple Choice Questions: Total	NS	NS	NS	-.39(.078)	NS
Test Total Score	NS	NS	NS	NS	NS

* Not significant ($\rho > .10$)
** Levels of significance are indicated in parentheses

on the pupils in their classrooms who are less able in reading. That effort might reflect in the skill growth of less able pupils, but would not necessarily raise the average classroom scores. An hypothesis to be tested would be whether the more diagnostically skilled teachers tend to work more with less able pupils.

The test of reading comprehension used was a norm-referenced measure of comprehension, the California Achievement Tests - Reading Comprehension subtest (CTB/McGraw-Hill, 1970 Edition). There were no relationships between second or fifth grade pupil skill growth on that test and scores from the major scoring categories on the film test. The criterion referenced decoding test, developed for the study, was not significantly related to teacher scores on the film test. The tests of reading achievement and application skills were also criterion referenced measures developed for the study. Poor pupil skill growth on the Reading Achievement Test at both grade levels was associated with good teacher scores on one (fifth grade teachers) or more (second grade teachers) of the major scoring categories of the film test. Poor second grade pupil skill growth on the Reading Application Test was related to good teacher scores on two of the major scoring categories of the film test.

The application and achievement tests were designed to include easy, medium and hard items which measured pupil skills on literal and inferential comprehension tasks. One-half of the items on each of the two tests were directed at literal comprehension and one-half were directed at inferential comprehension. It was previously pointed out that few of the teachers in the film test sample attended to those kinds of errors made by children in the film test. It was also

pointed out earlier that the good performance of second grade teachers on the film test was also consistently and negatively associated with observed positive pupil performance on comprehension tasks.

Each of the four videotaped children who read paragraphs was asked four questions after reading the paragraph; two were literal and two were inferential in nature. Any given child missed both literal questions and responded correctly to the two inferential questions, or missed both inferential questions and responded correctly to both literal questions. An hypothesis that could be tested would be whether teachers with highly developed diagnostic skills concentrate on the less sophisticated decoding skills in reading to the exclusion of the teaching of the more complex reading skills involved in literal nd inferential comprehension.

Summary and Conclusions

Teacher performance on the Diagnostic Film Test was analytically related to teacher performance on other reading and reading related tests, to observed teacher responses to pupils in their classrooms, and to observed and measured pupil reading performances. Some of the results were:

1) General Findings:

   a. Second grade teachers were more skilled at diagnosing problems and suggesting strategies for the remediation of pupil decoding problems in oral reading than were fifth grade teachers.

   b. Second and fifth grade teachers were about equally skilled at diagnosing problems and suggesting strategies for the remediation of pupil comprehension problems in oral reading.

   c. Neither second nor fifth grade teachers did particularly well in generating diagnoses and remediation strategies using the comparative standard of possible scores.

2) Performance on Other Reading and Reading Related Tests:

   a. The diagnostic and remediation skills of fifth grade teachers were related to their general knowledge about the teaching of reading.

   b. Teacher skills in the diagnosis and remediation of decoding problems were associated with teacher knowledge about the teaching of mathematics at both grade levels.

3) Observed Teacher Responses to Pupils During Reading Instruction:

   a. Second grade teachers who were more frequently observed helping individual pupils during reading were the more skilled diagnosticians.

   b. More highly developed diagnostic skills in second grade teachers were observationally associated with less of a tendency not to respond to individual pupils during reading activities.

4) Observed Pupil Performances in Reading:

a. Instances of pupil skills in phonics were observed more frequently in the classrooms of diagnostically skilled second grade teachers, but the same teacher skills were associated with poorer pupil comprehension skills.

b. Instances of positive pupil comprehension skills and general reading skills were observed more frequently in the classrooms of diagnostically skilled fifth grade teachers.

5) Measured Pupil Performances in Reading:

a. Second grade pupil skill growth in reading achievement and application tasks was negatively associated with teacher diagnostic skills.

b. There was no relationship between pupil skill growth on decoding tasks and the diagnostic skills of teachers.

All of the results should be regarded as hypotheses to be tested in future research. Teachers who took the test were enthusiastic about the Diagnostic Film Test, which they saw as measuring an important teaching skill in a relevant way. Probably because of the face validity of the film test, the comment, "This is the best test of all those we have taken," was not infrequent. Teachers, as well as the two professional staff members knowledgeable about reading* who administered the tests and sought teacher comments, felt that the instrument had potential as both a training and selection tool.

Those kinds of responses are perhaps adequate reasons in themselves for the further development of tests of this sort. But the

---

*The two were: Reginald Corder, Researcher and author of The Information Base for Reading - A Critical Review of the Information Base for Current Assumptions Regarding the Status of Instruction and Achievement in the United States (1971), research supported by the U.S. Department of Health, Education and Welfare; and Ann McCormick Piestrup, researcher and author in the field of reading.

relationships between the demonstration of teacher skills on the test, the observed evidence of pupil skills in related reading tasks, and the observed patterns of teacher responses to individual pupils all suggest that the Diagnostic Film Test is a test worthy of further research and development.

The relatively poor performance of both second and fifth grade teachers on the comprehension sections of the Diagnostic Film Test, the poor observed comprehension task performance of second grade pupils of diagnostically skilled teachers, and the inverse relationship of teacher skills demonstrated on the test and pupil achievement on literal and inferential comprehension tasks all suggest further exploration of the teaching priorities and skills of elementary school teachers.

References

Cooperative Primary Tests. Handbook. Educational Testing Service, 1967.

Cooperative Primary Tests. Reading Form 23A, Reading Form 23B, Word Analysis Form 13A, Word Analysis Form 13B. Educational Testing Service, 1965.

Corder, R. The Cooperative Primary Reading Test and its relationship to the California Miller-Unruh Program. Unpublished, private copy, 1970.

Goodman, K. S. Analysis of oral reading miscues: Applied psycho-linguistics. Reading Research Quarterly, 1969, 5, 1, 9-30.

Goodman, K. S. A linguistic study of cues and miscues in reading. Elementary English, 1965, 2, 639-643.

Gray, W. S. Gray Oral Reading Tests. Manual of Directions for Administering, Scoring, and Interpretation, Revised. Bobbs-Merrill Co., Inc., 1967.

Office of Compensatory Education. Word study skills for individualizing instruction. Book One and Book Two. Oakland Public Schools, 1973.

Sandoval, J. Decoding progression. Unpublished, private copy, 1973.

Schale, F. C. Changes in oral reading errors at elementary and secondary levels. Academic Therapy Quarterly, 1966, 1, 225-229.

Shankweiler, D. & Liberman, I. Misreading: A search for cues. In Kavanaugh, J. & Mattingly, I. G. The relationship between speech and reading. Cambridge, Massachusetts: MIT Press, 1972, 293-316.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill Book Company, 1956.

Weber, R. A linguistic analysis of first-grade reading errors. Reading Research Quarterly, 1970, 2, 427-431.

Weber, R. The study of oral reading errors: A survey of the literature. Reading Research Quarterly, 1968, 1, 96-119.

APPENDIX A

Reading Difficulties: Videotape Test of Teachers' Diagnostic Skills

READING DIFFICULTIES: VIDEOTAPE TEST OF TEACHERS' DIAGNOSTIC SKILLS

Name_____ Date_____

School_____ Grades_____

The materials read by each child are provided in this package. If you wish, you may use the list of words or paragraph provided to note the errors you hear each child make. Please complete the questions about each child's reading performance after each child reads.

SECTION ONE: CHILDREN READING WORD TESTS

Copyright © 1974 by Educational Testing Service. All rights reserved.

Child One: Jeffrey

His word list:

1. part	10. take
2. metal	11. dairy
3. buy	12. cake
4. right	13. gave
5. lives	14. about
6. squirt	15. faster
7. marble	16. where
8. now	17. just
9. hurt	18. upon

1. What is Jeffrey's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Jeffrey improve his reading skills?

3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?

   Materials:

   Human:

4. What resources would you need (if any) that are not presently available in your school or district?

   Materials:

   Human:

Child Two: Kelly

Her word list:

1. but	10. where
2. word	11. hat
3. leaf	12. them
4. eight	13. goes
5. please	14. because
6. sheet	15. much
7. on	16. clean
8. round	17. never
9. kind	18. would

1. What is Kelly's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Kelly improve her reading skills?

3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?

   Materials:

   Human:

4. What resources would you need (if any) that are not presently available in your school or district?

   Materials:

   Human:

-6-

Child Three: Karen

Her word list:

1. bells	10. foot
2. goats	11. but
3. boat	12. peach
4. which	13. always
5. opens	14. together
6. best	15. scout
7. rat	16. myself
8. there	17. every
9. laughs	18. use

1. What is Karen's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Karen improve her reading skills?

3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?

   Materials:

   Human:

4. What resources would you need (if any) that are not presently available in your school or district?

   Materials:

   Human:

-8-

Child Four: Chris

His word list:

1. draw	10. sunny
2. mile	11. shop
3. rid	12. hats
4. rope	13. found
5. one	14. us
6. core	15. point
7. won	16. before
8. think	17. light
9. throws	18. just

1. What is Chris' reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Chris improve his reading skills?

3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?

   Materials:

   Human:

4. What resources would you need (if any) that are not presently available in your school or district?

   Materials:

   Human:

SECTION TWO:
CHILDREN READING
PARAGRAPHS

-11-

Child Five: Andy

His paragraph:

RULES FOR TAKING OUT LIBRARY BOOKS

You cannot take out more than three books at one time. Take the card out of the back of each book you want. Write your name on each card, and leave the cards in the box on the library table. You may keep books from the blue shelves for one week and books from the white shelves for two weeks.

The questions he was asked:

- If Ruth kept a book from the blue shelf for one week, did she obey the library rules? (correct answer - yes)

- What is the best reason for returning books on time? (correct answer - so you will not lose them)

- Which shelf has books you may keep for two weeks? (correct answer - the white shelf)

- How many books may you take out at a time? (correct answer - three)

-12-

1. What is Andy's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would you use to help Andy improve his reading skills?

3. What resources (human or material) would you seek in your school or district to help you carry out a program for this pupil?

   Material:

   Human:

4. What resources woud you need (if any) that are not presently available in your school or district?

   Material:

   Human:

-14-

Child Six: Peter

His paragraph:

> There are many different kinds of whales. The blue whale is the biggest animal that ever lived, as far as anybody knows. It is so gigantic that it makes a large elephant look tiny. Even the huge dinosaurs would have looked small beside it. A baby blue whale weighs about 16,000 pounds. Not all whales are so big, though. One kind of whale that lives off in the rivers of India is only four feet long when it is fully grown.

The questions he was asked:

- How many different kinds of whales are there? (correct answer - the story doesn't tell)

- Does gigantic mean very large or fully grown? (correct answer - very large)

- From the story can you tell whether an elephant is smaller than some whales or bigger than the huge dinosaurs? (correct answer - smaller than some whales)

- What is the biggest animal that ever lived, as far as we know, a blue whale or a dinosaur? (correct answer - a blue whale)

1. What is Peter's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Peter improve his reading skills?

-16-

Child Seven: Julie

Her paragraph:

> Use a hard-boiled egg. Draw the eyes and mouth on the egg with wax crayons. Make them funny shapes. Next, dye the egg a solid color. Follow the directions on the dye box. Use bright colored paper for the nose and hat. Cut out a circle for the nose. Roll some paper into a tiny cone for the hat. Paste them both on the egg.

The questions she was asked:

- Is the nose made with a paper circle or a wax crayon? (correct answer - with a paper circle)

- Would you use crayons or dye first? (correct answer - crayons)

- Does the story tell how to make an egg bunny or an Easter clown? (correct answer - an Easter clown)

- Do the directions say to use a hard-boiled egg so the egg wouldn't break or so you could dye the egg? (correct answer - so the egg wouldn't break)

1. What is Julie's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Julie improve her reading skills?

-18-

Child Eight: Lisa

Her paragraph:

> Some nature museums lend animals to children to take home. Some of the animals are wild. The museums lend snakes, opossums, and racoons. Wild animals like these are taken home in locked cages. Wild animals may be kept one day. Tame animals such as rabbits and kittens may be kept a whole week. Rules come with the pets. The rules tell how to feed, clean, and handle the animals.

The questions she was asked:

- Does the story say that zoos or museums lend animals? (correct answer - museums)

- Would foxes or turtles come in a locked cage? (correct answer - foxes)

- Do rules come with the pets so the animal will get good care or so you will know if an animal is wild or tame? (correct answer - so the animals will get good care)

- May wild animals be kept one day or one week? (correct answer - one day)

1. What is Lisa's reading problem or problems? Be as specific as you can.

2. What strategies or approaches would <u>you</u> use to help Lisa improve her reading skills?

READING DIFFICULTIES:
VIDEOTAPE TEST OF
TEACHERS' DIAGNOSTIC SKILLS

Name_____ Date_____

School_____ Grades_____

SECTION THREE

copyright © ETS, 1974

-21-

Child Four: Chris

His word list:

1. draw	10. sunny
2. mile	11. shop
3. rid	12. hats
4. rope	13. found
5. one	14. us
6. core	15. point
7. won	16. before
8. think	17. light
9. throws	18. just

1. Chris' reading problem (or problems) include:
   (check all that apply)

   ____ initial consonant error(s)

   ____ medial consonant error(s)

   ____ final consonant error(s)

   ____ initial vowel error(s)

   ____ medial vowel error(s)

   ____ final vowel error(s)

   ____ reversal error(s)

   ____ basic sight vocabulary error(s)

   ____ consonant blend or digraph error(s)

   ____ vowel digraph or dipthong error(s)

2. Circle Chris' major problem or problems in the list above

-22-

Child Five: Andy

His paragraph:

RULES FOR TAKING OUT LIBRARY BOOKS

You cannot take out more than three books at one time. Take the card out of the back of each book you want. Write your name on each card, and leave the cards in the box on the library table. You may keep books from the blue shelves for one week and books from the white shelves for two weeks.

The questions he was asked:

- If Ruth kept a book from the blue shelf for one week, did she obey the library rules? (correct answer - yes)

- What is the best reason for returning books on time? (correct answer - so you will not lose them)

- Which shelf has books you may keep for two weeks? (correct answer - the white shelf)

- How many books may you take out at a time? (correct answer - three)

1. Andy's reading problem (or problems) include:
   (check all that apply)

   ____ hesitation

   ____ repetition

   ____ insertion of word(s) or letter(s)

   ____ omission of word(s) or letter(s)

   ____ decoding, gross pronunciation or punctuation error(s)

   ____ substitution of similarly shaped word(s) with different meaning(s)

   ____ substitution of word(s) that sound and look similar and have a similar meaning (or meanings)

   ____ interpretation or inference error(s)

   ____ literal meaning extraction error(s)

2. Circle Andy's major problem (or problems) in the list above.

APPENDIX B

Diagnostic Film Test
Scoring Manual

DIAGNOSTIC FILM TEST

SCORING MANUAL

Gail Hare

Patricia Elias

Copyright ©1975 by Educational Testing Service. All rights reserved.

SECTION I: DIAGNOSIS OF WORD READING ERRORS

Presentation of the Problem

In Section I of the Diagnostic Film Test, teachers are presented four videotape vignettes of children reading lists of words. Each list contains a total of 18 words. In each of the four cases, the child who is reading the list makes errors on nine of the 18 words. Six of the nine errors are of the same type. These six errors, grouped together, constitute what we will call the major or dominant error. The other three errors in any given list are of separate types and are termed minor, single occurrence errors.

Major or Dominant Errors

When a word is read, there are three general locations for a decoding error: the beginning of the word, the middle of the word, and the end of the word. The following four major or dominant error types were chosen for presentation:

1) Initial consonant error. (Example: "cake" is read by the child as "make".)
2) Final consonant error. (Example: "but is read by the child as "bun".)
3) Medial consonant error. (Example: "shared" read as "shamed".)
4) Medial vowel error. (Example: "sin" is read as "sun".)

Minor, Single Occurrence Errors

In any given sequence, two of the three minor errors are single examples of major errors. For example, if the major error in the list is with initial consonants, there will be single instances of a medial consonant or medial vowel error, and a final consonant error.

The third type of minor error is the same in each of the four cases. It is an example of a letter sequence reversal error. (For example, "rat" is read by the child as "tar".)

Scoring Generalizations

Teachers are awarded points for the diagnosis or identification of the major and minor errors, and also for presenting strategies for the correction of these major and minor errors. The general formula for scoring Section I will be outlined below. Examples will be presented of instances in which points were given and not given to teachers according to their answers. Answers given in common-sense or lay language are awarded as many points as answers phrased in educational jargon. However, in order to be awarded any points, answers must be problem specific. Global generalizations are not considered satisfactory or sufficient responses to the diagnosis problems presented. An answer can be awarded points no matter where it appears in the test protocol. For example, if the response given under "Diagnosis" is incomplete, but the problem is further explicated under "Strategy", diagnosis points are awarded for all information which appears under both questions.

The general scoring specifications are as follows:

Diagnosis of:	One major error = 4 points
	Three minor errors = 1 point each
	Possible total = 7 points

Strategies for Remediation: Major error = 2 points each, if clearly identified in relation to the major error

= 1 point each, if left general as to problem

(The number of strategies which may be presented to remediate the major problem is unlimited. The total points for strategies estimated below assumes that five strategies will be presented of two point value.)

minor errors = 1 point each

(Only one strategy for each minor error is awarded points.)

approximate total possible = 13 points

Each case thus has an approximate total point value of 20 points. Section I has an approximate total point value of 80 points (four cases each worth 20 points.)

Scoring the Diagnosis of the Major Error

The major error in the word reading cases involves decoding problems with consonants or vowels in various locations in a word. The individual cases for Section I are listed below, followed by the major error in each case:

Child One: Jeffrey - medial consonants or medial consonant blends

Child Two: Kelly - final consonants or consonant blends

Child Three: Karen - initial consonants or consonant blends

Child Four: Chris - medial vowels

Some basic generalizations cover the scoring of the major error. These are listed below, followed by some examples of the application of the scoring system.

1) Problem is defined by class of sound or letter (i.e., vowel or consonant) 1 point

2) Problem is defined by position in the word (i.e., initial, middle or medial, or final sound) 1 point

3) In addition to "1", "2", or "1 and 2" above, examples are given of the major errors (e.g., "substitutes P for T, b for h) 1 point

4) In addition to any of above, the problem is clearly stated as the major difficulty, or is implicitly stated as the major difficulty ("I would work most on ..."), or the number examples of the major error (six) is given. 1 point

Examples

"Jeffrey is missing the middle consonant."	2 points (Child One)
"Jeffrey does not recognize letters that appear within words."	1 point (Child One)
"Most of Kelly's consonant drill should be on ending sounds, such as in the words "but, word, hat, them, much."	4 points (Child Two)
"Kelly lacks self-confidence."	0 points (Child Two)
"Karen needs the most help on beginning consonant sounds."	3 points (Child Three)
"Karen guesses."	0 points (Child Three)
"Chris does not appear to recognize vowels or know their sounds."	1 point (Child Four)
"Medial vowel substitution is Chris' main problem."	3 points (Child Four)

Scoring the Diagnosis of the Minor Errors

The minor errors are similar for each of the three cases. As indicated earlier, there is only one example of each minor error. For each case, these are listed below:

Child One: Jeffrey - One reversal error; one final consonant error; one initial consonant error

Child Two: Kelly - One reversal error; one initial consonant error; one medial vowel error.

Child Three: Karen - One reversal error; one final consonant error; one medial vowel error.

Child Four: Chris - One reversal error; one final vowel error; one one initial consonant error.

There are three options for obtaining the one point awarded for the diagnosis of each minor error:

1) Problem is clearly stated in terms of class of sound involved (vowel or consonant) and location in the word (beginning, middle, end).

OR

2) Problem is generally stated as "trouble with beginning sound" or "consonant trouble", etc., and the example is given.

OR

3) The problem is generally stated, and it is clearly stated that there is only such problem on the list.

Scoring the Strategies Presented for Remediation of Major and Minor Errors

As indicated earlier, points are awarded for strategies presented in relation to the major error according to how specifically the strategy is identified with the problem. Points are not awarded if the errors have not been identified somewhere in protocol. Strategies for each of the minor

errors must be specific to those individual problems. Examples are given below of some acceptable and unacceptable strategy answers. Generalizations are not considered sufficient as strategies.

"Use sandpaper letters to help learn to identify letters like r, t, b, d"	Major error - 2 points Minor error - 1 point
"Individual skill lessons would help" (Error has been identified somewhere in protocol)	Major error - 1 point Minor error - 0 points
"I would see if Karen couldn't be given remedial reading help" or "Read words in sentences"	Major error - 0 points Minor error - 0 points

Section I Error Key

Child One: Jeffrey

His word list:

	(He reads)		(He reads)
1. part	pant	10. take	make
2. metal	medal	11. dairy	daisy
3. buy	but	12. cake	came
4. right		13. gave	
5. lives		14. about	
6. squirt	squint	15. faster	fatter
7. marble	ramble	16. where	
8. now		17. just	
9. hurt		18. upon	

Child Two: Kelly

Her word list:

	(She reads)		(She reads)
1. but	bus	10. where	there
2. word	work	11. hat	had
3. leaf	loaf	12. them	then
4. eight		13. goes	
5. please		14. because	
6. sheet	sheep	15. much	mush
7. on	no	16. clean	
8. round		17. never	
9. kind		18. would	

Child Three: Karen

Her word list:

	(She reads)		(She reads)
1. bells	tells	10. foot	food
2. goats	coats	11. but	put
3. boat	boot	12. peach	teach
4. which		13. always	
5. opens		14. together	
6. best	rest	15. scout	spout
7. rat	tar	16. myself	
8. there		17. every	
9. laughs		18. use	

Child Four: Chris

His word list:

	(He reads)		(He reads)
1. draw		10. sunny	funny
2. mile	mule	11. shop	sheep
3. rid	ride	12. hats	hits
4. rope	ripe	13. found	
5. one		14. us	
6. core	care	15. point	paint
7. won	now	16. before	
8. think		17. light	
9. throws		18. just	

SECTION II: DIAGNOSIS OF PARAGRAPH READING ERRORS

Presentation of the Problem

In Section II of the Diagnostic Film Test, there are four case presentations. In each case, a child reads a paragraph and answers four questions presented orally about what he has read. In each paragraph reading, seven errors are made. Each child also answers two of the four comprehension questions incorrectly. There are a total of nine reading comprehension errors for each of the four cases.

In paragraph reading, the most critical errors made, the errors which can most seriously affect reading comprehension, are errors of substitution and omission. These are the errors which are classified as the major or dominant errors in paragraph reading. There are four examples of the major error in each of the four paragraph reading cases. The other three reading errors are types which do not tend to seriously affect comprehension. These are the minor, single occurrence errors. The three minor errors are of the same type for each paragraph.

Major or Dominant Errors

In reading a paragraph, major errors are made by substitutions or omissions of letters and words that may or may not change the meaning of the paragraph and affect comprehension. These major errors can be broken down and grouped as follows:

Substitution errors

1) Phonologically similar words. The child substitutes words that sound the same. The words may or may not be real words, but they do not make sense in the context of the paragraph or sentence.

2) Graphically similar words. The child substitutes words that are similarly shaped. The words are often real words, but they do not make sense in context.

3) Phonologically and graphically similar words. The child substitutes words that sound and/or look similar. The words are real words that have a similar meaning to the word missed, and do make sense in context.

Omission Errors

4. Letters or words. The child omits critical letters or entire words, which are important to the comprehension of the paragraph.

The two errors each child makes on the comprehension questions following each paragraph are included in the major error scoring scheme. In each set of four questions, two of the questions test literal comprehension and two of the questions test inferential comprehension. A child either gets both inferential meaning questions correct and both literal meaning questions wrong, or vice versa. Among the four cases, two children miss both literal comprehension questions and two children make errors on both of the inferential comprehensions questions asked.

Minor Errors

The three types of single occurrence errors which appear in each paragraph reading are:

1) Hesitation: a pause or stop in the oral reading.

2) Repetition: a word is repeated in the course of reading the paragraph.

3) Insertion: an extra word is spoken in the course of reading the paragraph.

Specifications for Scoring Section II

Scoring Generalizations

As in Section I, teachers are awarded points for diagnosing and for presenting strategies to remediate the problems presented by each child. Though the scoring formula presented below is slightly more complex than that for Section I, it follows the same rationale and utilizes the same major/minor error problem construction. Since there are slight differences, the major error scoring formula is presented separately for each case. As in Section I, diagnosis answers termed in standard English are as acceptable as those phrased in jargon. Global generalizations are not adequate. The specific problem must be dealt with.

The possible points are the same as for Section I, and those general scoring specifications are repeated here.

Diagnosis of:	One major error = 4 points
	Three minor errors = 1 point each
	Possible total = 7 points

Strategies for Remediation: Major error = 2 points each, if clearly identified in relation to the major error

= 1 point each, if left general as to problem

(The number of strategies which may be presented to remediate the major problem is unlimited. The total points for strategies estimated below assumes that five strategies will be presented of two point value.)

minor errors = 1 point each

(Only one strategy for each minor error is awarded points.)

approximate total = 13 points

The approximate total value per case is 20 points. The approximate total point value for Section II is 80 points.

Scoring the Diagnosis of the Major Error - Child Five: Andy

The major error made by this child in reading his paragraph was the substitution of phonologically similar words that change meaning. The scoring formula is as follows:

1) Problem is explained as "incorrect words", "missing words", "miscalls words", "substitutes words", "poor skills in sounding words". 1 point

2) Problem is identified generally as 1) above, and the errors made are listed. 1 point

3) Problem is identified as 1) above and it is stated that the mispronunciation throws off comprehension, changes meaning, etc. OR

The problem is further identified as involving words that sound alike. 1 point

4) It is indicated that the errors described are the major problem.

OR

It is indicated that the errors the child made caused him to miss the literal comprehension questions (synonyms accepted). 1 point

Examples

"Andy mostly substitutes words that sound alike, code for card, wait for while." 4 points

"Andy did not read carefully." 0 points

Scoring the Diagnosis of the Major Error - Child Six: Peter

The major error made by this child was the substitution of graphically similar words with different meanings. This case is scored as follows:

1) Problem is described as "incorrect words", "missing words" "miscalls words", "substitutes words", "poor skills in sounding words". 1 point

2) Problem is described as 1) above, and the actual errors made are listed. 1 point

3) Problem is further identified as substitution of similarly shaped words.

OR

It is stated that the mispronunciation alters meaning. 1 point

4) It is indicated that the problem described is the major problem.

OR

It is indicated that the problem described caused the child to miss the inferential comprehension questions (synonyms accepted). 1 point

Examples

"Peter misread through for though, where for when." 2 points

"Peter makes mistakes with words that look alike, and his mistakes on the words cause him to misunderstand the point of the story." 3 points

Scoring the Diagnosis of the Major Error - Child Seven: Julie

Julie's major problem is the substitution of phonologically and graphically similar words that also have similar meanings. Her major error is scored as follows:

1) Problem is stated as "substituting words", "incorrect words", "missing words", miscalls words", "poor skills in sounding words". 1 point

2) Problem is described as 1) above, and the errors made are listed. 1 point

3) Problem is further identified as a substitution of words that sound and look similar.

OR

It is stated that the substitutions do not affect the meaning. 1 point

4) It is indicated that the errors described are the major problem.

OR

It is indicated that the problem caused the child to make errors on inferential comprehension problems (synonyms accepted). 1 point

Examples

"She misreads some words." 1 point

"Julie substitutes words, but retains the sentence meaning, for instance, she reads crayons - crayolas." 3 points

Scoring the Diagnosis of the Major Error - Child Eight: Lisa

Lisa's major problem is the omission of words and letters which are important to the meaning of the paragraph, the final "s" in particular. The diagnosis of this major error is as follows:

1) Problem is stated as omission of words. 1 point

2) Problem is stated as omission of critical letters or omission of the final "s". 1 point

3) Problem is described as 1), 2) or 1) and 2) above and the errors made are listed.

OR

It is stated that the ommissions affect the meaning or comprehension of the paragraph. 1 point

4) It is stated that the errors described are the major problem.

OR

It is indicated that the problem caused the child to make literal comprehension errors (synonyms accepted). 1 point

Examples

"Lisa omitted the final "s" sound, and also some words." 2 points

"Lisa did not pay attention to what she was reading." 0 points

Scoring the Diagnosis of the Minor Errors

Identification of the three minor errors is given one point for each error identified. The three single occurrence errors are the same for each paragraph, and identification should be made in a reasonably straightforward manner either by stating generally that hesitation, repetition and insertion (synonyms for these are acceptable) have occurred, or by citing the exact example from the paragraph.

Scoring the Strategies Presented for Remediation of Major and Minor Errors

As in Section I, points are awarded for strategies presented in relation to the major error according to how specifically the strategy is identified with the problem. Strategies for the minor errors must be specific to those individual problems. Below are examples of strategy answers and the points awarded them.

"Let him read aloud and guide his accuracy in word meaning in different lessons, such as science, social studies."	Major error - 1 point
"Needs some help on sight vocabulary, so I would use flashcards to strengthen recognition of words like though, when, etc."	Major error - 2 points
"I would make her pay attention to what she was reading."	Major error - 0 points
"This child should be exposed to high interest material which he can read aloud to an aide. This will reduce his hesitation."	Minor error - 1 point

Section II Error Key

Child Five: Andy

His paragraph:

RULES FOR TAKING OUT LIBRARY BOOKS

You cannot take out more than three books at one time. Take the card out of the back of each book you want. Write your name on each card, and leave the cards in ▲ the box ▲ on the library table. You may keep books from the blue shelves for one week [] and books from the white shelves for two weeks.

(He reads) there cade ▲ in ... ▲ of ▲ libery [] (pause) wait

The questions he was asked:

- If Ruth kept a book from the blue shelf for one week, did she obey the library rules? (correct answer - yes)

- What is the best reason for returning books on time? (correct answer - so you will not lose them)

- Which shelf has books you may keep for two weeks? (correct answer - the white shelf)

"the blue shelf" (literal)

- How many books may you take out at a time? (correct answer - three)

"two" (literal)

Child Six: Peter

His paragraph:

There are many different kinds of whales. The blue whale is the biggest animal that ever lived, as <u>far</u> as anybody knows. It is so gigantic that it makes a large elephant look <u>tiny</u>. Even the huge dinosaurs would have▲looked small beside it. A baby blue whale weighs about▯16,000 pounds. Not all whales are so big, <u>though</u>. One kind▲of whale that lives off in the rivers of India is only four feet long <u>when</u> it is fully grown.

(He reads)

for

tall

▲ of

▯ (pause)

through ▲one kin

where

The questions he was asked:

- How many different kinds of whales are there? (correct answer - the story doesn't tell)

- Does gigantic mean very large or fully grown? (correct answer - very large)

"fully grown" (inference)

- From the story can you tell whether an elephant is smaller than some whales or bigger than the huge dinosaurs? (correct answer - smaller than some whales)

"bigger than some dinosaurs" (inference)

- What is the biggest animal that ever lived, as far as we know, a blue whale or a dinosaur? (correct answer - a blue whale)

Child Seven: Julie

Her paragraph:

Use a hard-boiled egg. Draw the eyes and mouth on the egg with wax crayons. Make them funny shapes. Next, dye the egg a solid color. Follow the directions on the dye box. Use bright colored paper for the nose and hat. Cut out a circle for the nose. Roll some paper into a tiny cone for the hat. Paste them both on the egg.

(She reads) broiled crayolas colors of a a (pause)

The questions she was asked:

- Is the nose made with a paper circle or a wax crayon? (correct answer - with a paper circle)

- Would you use crayons or dye first? (correct answer - crayons)

- Does the story tell how to make an egg bunny or an Easter clown? (correct answer - an Easter clown)

"an egg bunny" (inference)

- Do the directions say to use a hard-boiled egg so the egg wouldn't break or so you could dye the egg? (correct answer - so the egg wouldn't break)

"so you could dye the egg" (inference)

Child Eight: Lisa

Her paragraph:

(She reads)

Some nature museums ~~lend~~ animals to children to take home. Some of the animals are wild. The museums lend [] snakes, opossums, and racoons. Wild animals like these are taken home in locked cages. Wild animals may be kept one day. ▲ Tame animals such as rabbits and kittens may be kept a whole week. Rules come with the pets. The rules ~~tell~~ how to ▲ feed, clean, and handle the animals.

— (omits)
[] (pause)
/ (omits)
▲ only
▲ come — (omits)
/ (omits)

The questions she was aske

- Does the story say that zoos or museums lend animals? (correct answer - museums)

"zoos" (literal)

- Would foxes or turtles come in a locked cage? (correct answer - foxes)

- Do rules come with the pets so the animal will get good care or so you will know if an animal is wild or tame? (correct answer - so the animals will get good care)

- May wild animals be kept one day or one week? (correct answer - one day)

"one week" (literal)

SECTION III: DIAGNOSIS OF ORAL READING ERRORS IN MULTIPLE-CHOICE FORMAT

In Section III, one case is presented again from Section I, and one from Section II. The videotape cases involved are Child Four: Chris and Child Five: Andy. Teachers are asked to choose from a list of possible reading errors those errors which apply to the specific cases presented. In addition to checking all errors which applied, the major error or problem was to be circled.

In each of the two cases presented, there are two options on the list of problems which relate to the major error. There are three options presented which relate to the minor error. The scoring of this section is as follows:

Five error options correctly checked = 5 points

Two major error options correctly circled = 2 points

Possible total for each case presented = 7 points

Possible total for Section III = 14 points

Points are deducted as follows for incorrect answers:

Incorrectly checked error = -1/2 point

Incorrectly checked error, circled as a major error = -1 point

Minor error incorrectly circled as a major error = -1/2 point

The key for scoring the two cases in section III are presented on the following pages.

Section III Correct Answer Key

Child Four: Chris

1. Chris' reading problem (or problems) include:
(check all that apply)

✓ initial consonant error(s)

___ medial consonant error(s)

___ final consonant error(s)

___ initial vowel error(s)

✓ medial vowel error(s)

✓ final vowel error(s)

✓ reversal error(s)

___ basic sight vocabulary error(s)

___ consonant blend or digraph error(s)

✓ vowel digraph or dipthong error(s)

2. Circle Chris' major problem or problems in the list above

Child Five: Andy

1. Andy's reading problem (or problems) include:
(check all that apply)

- ✓ hesitation
- ✓ repetition
- ✓ insertion of word(s) or letter(s)
- ___ omission of word(s) or letter(s)
- ✓ decoding, gross pronunciation or punctuation error(s)
- ___ substitution of similarly shaped word(s) with different meaning(s)
- ___ substitution of word(s) that sound and look similar and have a similar meaning (or meanings)
- ___ interpretation or inference error(s)
- ✓ literal meaning extraction error(s)

2. Circle Andy's major problem (or problems) in the list above.